Remarks

The present amendment responds to the final Official Action dated April 4, 2005. The Official Action rejected claims 1-19 under 35 U.S.C. §112, second paragraph as being indefinite for purportedly using confusing language in claim 1. Claim 1 was rejected under 35 U.S.C. §102(e) based on Gusack U.S. Patent No. 6,356,897 (Gusack). Claims 2-19 were objected to as being dependent upon a rejected base claim, but were indicated to be allowable if rewritten to overcome the §112, second paragraph rejection and to include all of the limitations of the base claim. These grounds of rejection are addressed below.

Claim1 has been amended to be more clear and distinct. This amendment, as described below, makes explicit what is clearly implicit in the claims prior to this amendment and addresses the Examiner's Section 112 rejection. Claims 1-19 are presently pending.

Section 112, Second Paragraph Rejection

In response to this rejection, claim 1 has been amended to clarify that the method addresses the problem of reformulating raw data appearing in a delineated table region of an electronic document into a table structure. Claim 1, as presently amended, reads as follows:

> 1. (currently amended): A method for reformulating raw data appearing in a delineated table region of an electronic document into a table structure, comprising the steps of:
>
> a) reading raw data, said raw data spatially arranged in a delineated table region of an electronic document, said raw data, as read, lacking hierarchical arrangement sufficient to enable a logical query of said raw data based on said spatial arrangement;
>
> b) creating a binary tree using a hierarchical clustering of a plurality of words included in said raw data;

c) segregating a plurality of columns from the raw data using a breadth-first traversal algorithm;

d) identifying column headers, if any, from the plurality of columns using a first heuristic algorithm;

e) identifying row headers, if any, from the column headers using a second heuristic algorithm;

f) segregating at least one row from the raw data using a row determination algorithm; and

g) storing the plurality of columns and the at least one row into a table structure.

By reformulating raw data to a table structure, the claimed invention advantageously provides a table structure where a subsequent logical query may be used to retrieve data from the table structure. The Official Action appears to misconstrue the term "logical query" when it equates text scanning as an example of a logical query. A logical query is a form of a query that involves selecting a table entry whose attributes meet specific logical criteria. Referring to Fig. 1 of the present specification, for example, a logical query may select or retrieve the company that has the minimum opening price. In this example, Red, Inc would be returned because its opening price, 22 ¼, is the lowest of all the companies in the table. Text scanning, on the other hand, merely returns a string of characters that match the string of characters used in the query. Consequently, in a text scanning system, a query would have to be constructed with the text "22 ¼" specified. Because raw data appearing in a delineated table region of an electronic document such as a displayed table in a Word® document has no hierarchical arrangement, a logical query on raw data cannot be performed.

Claim 1 has been amended to clarify that the method reformulates raw data into a table structure. In particular, the method begins with raw data lacking hierarchical arrangement

sufficient to enable a logical query of said raw data based on said spatial arrangement and, through the recited steps, ends with storing intermediate output into a table structure. Support for the term "raw data" can be found, for example, at page 6, lines 8-17. Replacing the term "tabled data" with "raw data" to make clear the distinction between raw data lacking hierarchical arrangement that is read in step a) and the table structure stored upon execution of steps b) –g) which reformulates the raw data into a table structure enabling logical queries. Step g) of claim 1 has been introduced to complete the reformulation into a table structure which is stored and which can subsequently be logically queried. Support for this amendment can be found, for example, at page 16, lines 17-18. This rejection should now be overcome.

The Art Rejections

Claim 1 was rejected under 35 U.S.C. §102(e) based on Gusack. Applicants respectfully traverse this rejection and request that it now be withdrawn, in view of the above amendments and the discussion herein.

Claim 1 has been amended to recite, as its first step, that raw data are read as input. The raw data are spatially arranged in a delineated table region of an electronic document. However, the raw data, as so read, lack hierarchical arrangement sufficient to enable a logical query of the raw data based on said spatial arrangement. The preamble of claim 1, as amended, recites a method for reformulating raw data appearing in a delineated table region of an electronic document into a table structure.

The raw data as recited in the first step of claim 1 are electronic, but they are not stored in the form of a table structure, or otherwise in an array of fields that can be logically queried as

described above. According to the claimed invention, such raw data appearing in a delineated table region of an electronic document are subjected to a series of steps that reformulates such raw data into a table structure that can be logically queried. This method, for example, addresses the need for such reformulations as applied to such raw data which are generated by various software applications such that they may or may not contain border lines, a fixed number of blank spaces between columns, multi-line rows, multi-line column headers, or a clearly vertical column definition due to skewing, for example. In such exemplary cases, as well as many others, the present method can be advantageously applied for reformulating raw data into a table structure and enabling structured queries in raw data appearing in a delineated table region of an electronic document.

Gusack fails to disclose and fails to suggest a method step of reading raw data, said raw data spatially arranged in a delineated table region of an electronic document, said raw data, as read, lacking hierarchical arrangement sufficient to enable a logical query of said raw data based on said spatial arrangement. To the contrary, Gusack discloses database structures containing data and subjected to various linking and query procedures. Gusack fails to disclose and fails to suggest any method for reformulating raw data that are not stored in a prior-generated structured database that can be queried.

Gusack is entitled "Associative Database Model for Electronic-Based Informational Assemblies". Gusack summarizes an object of providing relational database structures as follows:

> Accordingly, it is an object of this invention to provide a unique indexing system for an assembly of informational items stored on electronic-based media, constructed from at least one registration table and set of program instructions that assigns at least one unique alphanumeric indicium to each data table included in the database structure and, therefore provides a means for calculating at least one unique domain of alphanumeric indicia for each table registered in the data set, in turn, providing a means for calculating at least one unique indicia for each record and each field defined within each record in the entire data set, and therefore providing a higher degree of order, integrity, continuity, and user convenience in accessing and creating knowledge from the assembly of informational items stored in said data tables than data structures of prior art. Gusack, col. 3, lines 43-57; see also, col. 3, lines 30-36.

Thus, it is seen that Gusack starts with a database structure, to which data are added and then manipulated. The Official Action makes reference to Fig. 9 of Gusack and its disclosure at col. 15, lines 27-53. Fig. 9 and the cited passage of Gusack disclose relational database structures. These citations are part of Gusack's disclosure of an embodiment of an indexing system, beginning at col. 11, line 51. The indexing system assigns to each record in a plurality of data tables, an alphanumeric indicium that is unique for all records stored in the data set. Col. 11, lines 51-54. Referring to Fig. 6 of Gusack, the first data table (601) is a table registry containing a TIN field (603) for storing a unique alphanumeric indicium. Gusack, col. 12, lines 6-10. Hence this data table, with which Gusack's indexing system begins, is itself a relational database, capable of assigning and storing Gusack's unique alphanumeric indicium. Referring to Fig. 6, each of the other four data tables 619, 625, 631 and 637 includes an L# field 621, 627, 633 and 641, respectively, for storing unique record identifiers. Gusack, col. 12, lines 55-57. Gusack's five data tables shown in Fig. 6 clearly do not lack hierarchical arrangement sufficient to enable a logical query of the raw data based on its spatial arrangement. The unique record identification

system ("URIS") is then generated by this indexing. Col. 13, lines 38-40. Once the data set is stored to provide the unique record identification indicia, a large number of options become available to structure relational links between records stored in tables included in the data set. Col. 13, lines 57-62. The actual links thus generated are then removed from each record and located in a separate set of specialized linking tables, constituted by the central linking table system ("CLTS"). Gusack, col. 14, lines 3-24. The Gusack passage on col. 15 cited by the Official Action states, "[t]he CLTS and URIS work together to allow a user to view linked records in said user interface." This discussion and Fig. 9 to which it refers, address a relational database.

Gusack accordingly fails to disclose and fails to suggest, at col. 15, in Fig. 9 or elsewhere, "reading raw data, said raw data spatially arranged in a delineated table region of an electronic document, said raw data, as read, lacking hierarchical arrangement sufficient to enable a logical query of said raw data based on said spatial arrangement," as presently claimed in claim 1. Applicants therefore respectfully traverse the assertions on page 2 and 3 of the Official Action that Gusack discloses a binary tree, a row determination algorithm, a breadth first traversal algorithm, and first and second heuristic algorithms, all as defined in claim 1.

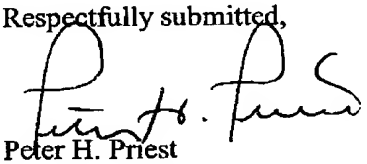
Gusack further fails to disclose and fails to suggest a step of "creating a binary tree from such tabled data using a hierarchical clustering of a plurality of words included in said raw data," as presently claimed in claim 1. Gusack additionally fails to disclose and fails to suggest a step of "segregating a plurality of columns from the raw data using a breadth-first traversal algorithm," as presently claimed in claim 1. Gusack additionally fails to disclose and fails to

suggest a step of "identifying column headers, if any, from the plurality of columns using a first heuristic algorithm," as presently claimed in claim 1. Gusack further fails to disclose and fails to suggest a step of "identifying row headers, if any, from the column headers using a second heuristic algorithm," as presently claimed in claim 1. Gusack additionally fails to disclose and fails to suggest a step of "segregating at least one row from the raw data using a row determination algorithm," as presently claimed in claim 1. Finally, Gusack additionally fails to disclose and fails to suggest a step of "storing the plurality of columns and the at least one row into a table structure," as presently claimed in claim 1.

Conclusion

All of the presently pending claims, as amended, appearing to define over the applied references, withdrawal of the present rejection and prompt allowance are requested.

Respectfully submitted,



Peter H. Priest
Reg. No. 30,210
Priest & Goldstein, PLLC
5015 Southpark Drive, Suite 230
Durham, NC 27713-7736
(919) 806-1600